

# HT1220 Report

## RECORD INFORMATION

Gene ID: 1220  
Sequence ID: 1220  
Protein ID: 1220  
Sequence name: thrombospondin 1, alt. transcript 1  
Genome: nucleus  
Taxon: Homo sapiens  
Locus: 1220  
Common Name: thrombospondin 1  
Role ID: 40

Coding sequence length: 3513 nt  
Transcript sequence length: 5722 nt  
Expression data: THC201673

## ACCESSION DATA

HT1220 is derived from accessions(s):

SP:P07996 (THROMBOSPONDIN 1 PRECURSOR.)  
GB:X04665 (Human mRNA for thrombospondin)  
GB:X14787 (Human mRNA for thrombospondin)  
GB:U12471 (thrombospondin-p50 {Homo sapiens})  
GB:M99425 (Human thrombospondin mRNA, 3' end.)  
PIR:G01478 (thrombospondin-p50 - human (fragment))  
GB:U12471 (Human thrombospondin-1 gene, partial cds.)  
GB:J04835 (Human thrombospondin gene, exons 1, 2 and 3.)  
GB:M25631 (Homo sapiens (clone lambda-TS-33) thrombospondin (THBS) mRNA, 5' end.)

## ALTERNATIVE SPLICE INFORMATION

Alternative splice forms for this gene:

HT3987 thrombospondin 1, alt. transcript 2

## MAPPING DATA

GDB accession(s) for this gene:

GDB ID:      Symbol  
-----

FIGURE 1A

## cDNA FEATURES

Feature	End 5	End 3
coding_seq	112	3624
3'UT	3625	5722
spjunc_h	1235	1236

## SEQUENCE

nucleotide:

ggacgcacaggaattccccgcgccccctccagccctcgccgcctcgccaccgctcccggc  
 cgccgcgctccgggtacacacaggatccctgctgggcaccaacagctccaccatggggctg  
 gcctggggactagggctcctgttccctgatgcatgtgtgtggcaccaaccgcatccagag  
 tctggcggagacaacagcgtgtttgacatccttgaactcacccggggccgcccgcgaagggg  
 tctgggcgcgactgggtgaagggccccgaccctccagcccagctttccgcatcgaggat  
 gccaacctgatccccctgtgcctgatgacaagtccaagacctgggtggatgctgtgcgg  
 gcagaaaagggtttcctccttctggcatccctgaggcagatgaagaagaccggggcacg  
 ctgctggccctggagcggaaaagaccaactctggccaggctcttcagcgtgggtgtccaatggc  
 aaggcggggcacccctggacctcagcctgaccgtccaaggaaagcagcacgtgggtgtctgtg  
 gaagaagctctcctggcaaccggccagtggaaagagcatcacctgtttgtgcaggaagac  
 agggcccagctgtacatcgactgtgaaaagatggagaatgctgagttggacgtcccatc  
 caaagcgtcttcaccagagacctggccagcatcgccagactccgcatcgcaaaggggggc  
 gtcaatgacaatttcagggggtgtctgcagaatgtgaggtttgtctttggaaccacacca  
 gaagacatctcaggaacaaggctgtccagctctaccagtgctcctcctcaccttgac  
 aacaacgtgggtgaatgggtccagccctgccatccgcactaacctacattggccacaagaca  
 aaggacttgcaagccatctgcggcatctcctgtgatgagctgtccagcatgggtcctggaa  
 ctcaggggctgcgcaccattgtgaccacgctgcaggacagcatccgcaaagtgaactgaa  
 gagaacaaagagtggcccaatgagctgaggcggcctcccctatgctatcacacggagtt  
 cagtacagaaaataacgaggaatggactgttgatagctgcactgagtgctcactgtcagaac  
 tcagttaccatctgcaaaaagggtgtcctgccccatcatgocctgctccaatgcccagtt  
 cctgatgggaatgtgtcctcgtgttggtggccagcagctctgcggacgatgggtgggtct  
 ccattggtccgagtggaacctcctgttctacgagctgtggcaatgggaattcagcagcgcgcc  
 cgctcctgcgatagcctcaacaaccgatgtgagggctcctcggtccagacacggacctgc  
 cacattcaggagtggtgacaaaagatttaaacaggatgggtggctggagccactgggtccccg  
 tgggtcatcttgttctgtgacatgtgggtgatgggtgtgatccacaaggatccggctctgcaac  
 tctcccagccccagatgaatgggaaaccctgtgaaggcgaagcgcgggagaccaaagcc  
 tgcaagaaagacgcctgccccatcaatggaggctggggctcctgggtcaccatgggacatc  
 tgttctgtcacctgtggaggagggtacagaaacgtagtcgtctctgcaacaaccccgca  
 cccagtttgaggagcaaggactgcgttgggtgatgtaacagaaaaccagatctgcaacaag  
 caggactgtccaattgatggatgcctgtccaatccctgctttgocggcgtgaagtgtact  
 agctaccctgatggcagctggaaaatgtgggtgcttgtccccctgggttacagtggaaaatggc  
 atccagtgcacagatgttgatgagtgcaagaagtgcctgatgcctgcttcaaccacaat  
 ggagagcaccgggtgtgagaacacggaccgccggtacaactgcctgcccctgccccccacgc  
 ttacccggtccacagcccttcgggccagggtgtcgaaatgccacgggccaacaaacagggtg  
 tgcangccccgtaacccctgacagcggatgggacccacgactgcaacaagaacgccaagtgc  
 aactacctgggcccactatagcgacccccatgtaccgctgcagtgcaagcctggctacgct  
 ggcaatggcatctctgcggggaggacacagacctggatgggtggcccaatgagaacctg  
 gtgtgcgtggccaatgcgacttaccactgcaaaaaggataattgccccaaccttcccaac  
 tcagggcaggaagactatgacaaggatgggaattgggtgatgcctgtgatgatgacgatgac  
 aatgataaaattccagatgacaggggacaactgtccattccattacaaccagctcagtat  
 gactatgacagagatgatgtgggagaccgctgtgacaactgtccctacaaccacaacca

FIGURE 1B

gatcaggcagacacagacaacaatggggaaggagacgcctgtgctgcagacattgatgga  
gacgggtatcctcaatgaacgggacaactgccagtagctctacaatgtggaccagagagac  
actgatatggatgggggttggagatcagtgtagacaattgcccccttggaaacacaatccggat  
cagctggactctgactcagaccgcattggagatacctgtgacnacaatcaggatattgat  
gaagatggccaccagaacaactctggacaactgtccctatgtgcccattgccaaaccaggct  
gaccatgacaaagatggcaaggagatgcctgtgaccacgatgatgacaacgatggcatt  
cctgatgacaaggacaactgcagactcgtgccccaatcccgaccagaaggactctgacgyc  
gatggctcgaggatgcctgcaaagatgattttgaccatgacagtggtgccagacatcgat  
gacatctgtcctgagaatgttgacatcagtgagaccgatttccgcgattccagatgatt  
cctctggacccccaaaggacatccccaaatgaccctaactgggttgtagccatcagggt  
aaagaactcgtccagactgtcaactgtgatcctggactcgtgtagggttatgatgagttt  
aatgctgtggacttcagtggcaccttcttcatcaacaccgaaaggagcagatgactatgct  
ggatttgtcttggctaccagtcacgcgcgcttttatgttgtgatgtggaagcaagtc  
accagtcctacttgggacaccaacccccacgagggtcagggatactcgggcctttctgtg  
aaagtgttaactccaccacagggcctggcgagcacctgcggaacgcctgtggcacaca  
ggaaacacccctggccagggtgcgcacccctgtggcatgaccctcgtcacataggctggaa  
gatttcacgcctacagatggcgtctcagccacaggccaaagacgggtttcattagagtg  
gtgatgtatgaagggaagaaaatcatggctgactcaggacccatctatgataaaacctat  
gctgggtgtagactagggttgtttgtcttctctcaagaaatgggtgttcttctctgacctg  
aaatagcaatgtagagatccctaactcatcaaattgttgattgaaagactgatcataaacc  
aatgctgggtattgcaccttcttggaaactatgggcttgagaaaacccccaggatcacttctc  
cttggttctcttcttcttctgtgcttgcatcagtggtgactcctagaacgtgacgacctgc  
ctaaagaaaatgcagtttctcaaaaacagactcatcagcattcagcctccaatgaataagac  
atcttccaagcatataaacaattgcttttggtttctcttttgaaaaagcatctacttgcctc  
agttgggaagggtgccattccactctgcctttgtcacagagcagggtgctattgtgaggc  
catctctgagcagtgaggactcaaaagcattttcaggcatgtcagagaaggaggactcact  
agaattagcaaacaaaaccacctgacatcctccttcagggaacacggggagcagaggcca  
aagcactaaggggagggcgcataccagagacgattgtatgaagaaaatattggaggaaactg  
ttacatgttccgtactaagtcattttcaggggattgaaagactattgctggatttcatga  
tgctgactggcggttagctgatttaacccatgtaaataggcacttaaatagaagcaggaaag  
ggagacaaagactggctcttggaacttctccttgatccccacccttactcatcacttgc  
agtggccagaattagggaatcagaatcaanccagtgttaaggcagtgctggctgccattgc  
ctggtcacattgaaattgggtggcttcatcttagatgtagcttgtgcagatgtagcaggaa  
aataggaaaaacctaccatctcagtgagcaccagctgcctcccaaaggaggggcagccgtg  
cttatattttttatgggttacaatggcacaaaatttatatcaacctaaactaaaacattcctt  
ttctcttttttccgtaattactaggtagttttctaattctctcttttggaaagtatgattt  
ttttaaagtctttacgatgtaaaaatatttatttttacttattctggaagatctggctga  
aggattattcatggaacaggaagaagcgtaaagactatccatgtcatctttgttgagagt  
cttcgtgactgttaagattgttaaatacagattattttattaactctgttctgctggaaatt  
taggcttcatacgggaagtgtttgagagcaagtagttgacatttatcagcaaatctcttg  
caagaacagcacaaaggaaaatcagtcataaagctgctctgcccccttgtgctcagagtgg  
atggttatgggattccttttttctctgttttatcttttcaagtggaaattagttgggtatcc  
atttgcaaatgttttaaatgcaaaagaaagccatgaggtcttcaatactgttttacccca  
tcccttgtgcatatttccaggggagaaggaaagcatatacacttttttcttcttcttcttcc  
aaaagagaaaaaaatgacaaaagggtgaaacttacatacaaatattacctcatttgttggtg  
tgactgagttaaagaaattttggatcaagcggaaagagtttaagtgtctaacaaacttaaa  
gctactgtagtacctaaaaggcagtggttgtagatagcataaaaactctgcagagaagta  
ttcccaataaggaaatagcattgaaatgttaaatataaatttctgaaagtattgtttttt  
tctatcatctgggtataccattgctttatttttataaattattttctcattgccattggaa  
tagaatattcagattgtgtagatatgctattttaataatttatcaggaaatactgctgtg  
agagtttagtattttctatttttatataatgtttgcacactgaattgaagaattgttggtt  
tttcttttttttgtttttttttttttttttttttttgtctttgacctccatttttta  
ctatttgcaataaccttttcttaggaatgtgcttttttttgtacacatttttatccattt  
cacattctaaagcagtgtaangtgtatattactgtttcttatgtacaaggaaacaacata  
aatcatatggaaatttatattt

protein:

MGLAWGLGVLFLMHVCGTNRI PESGGDNSVFDI FELTGAARKGSORRLVKGPDPSSPAFR

FIGURE 1C

IEDANLIPPVPDDKFQDLVDAVRAEKGFLLLASLRQMKKTRGTLALERKDHSQQVFSVV  
SNGKAGTLOLSLTVQKGQHVVSVEEALLATGQWKSITLQVQEDRAQLYIDCEKMEAELO  
VPIQSVFTRDLASIAARLRIAGGVNDNFQGVLOQNVRFVFGTTPEDILRNKGCSSSTSVLL  
TLONNVVNGSSPAIRTNVYIGHKTKDLQAI CGISCELSMVLELRGLRTIVTTLQDSIRK  
VTBENKELANELRRPPLCYHNOVQYRNNEBWTVDSCTECHCQNSVTICKKVSCPINPCSN  
ATVPDGECCPRCWPSDSADDGWSPWSEWTSCSTSCGNGIQQRGRSCDSLNNRCEGSSVQT  
RTCHIQECDKRFKQDGGWSHWPWSSCSVTGCGDVITRIRLCNSPSPQMNGKPCEGEARE  
TKACKKDACPINGGWGPWSPWDICSVTCGGGVQKRSRLCNPAPQFGGKDCVGDVTENQI  
CNKQDCPIDGCLSNPCFAGVKCTSYPDGSKWCGACPPGYSGNGIQCTDVDECKEVPDACF  
NHNGEHRCENTDPGYNCLPCPPRFTGSQPFQGVQVEHATANKQVCKPRNPCTDGTHDCNKN  
AKCNYLGHYSDPMYRCECKPGYAGNGIICGEDTDLDGWPENLVCVANATYHCKKDNCPN  
LPNSGQEDYDKDGI GDACDDDDNDKI PDDRDNCPFHYNPAQYDYDRDDVGDRCDNCPYN  
HNPDAQADTDNNGEGDACAADIDGDGILNERDNCQYVYNVDQRDTDMDGVDGQCDNCPLEH  
NPDQLSDSDRIGDTCNNQDIDEDGHQNNLDNCPYVPANQADHDKDGKGDACDHDDDN  
DGIPDDKDNCRCLVPNPDQKSDGDGRGDACKDDFDHDSVPDIDDICPENVDISETDFRRF  
QMIPLDPKGTSQNDPNWVVRHQGKELVQTVNCDPGLAVGYDEFNAVDFSGTFFINTERDD  
DYAGFVFGYQSSSRFYVVMWKQVTQSYWDTNPTRAQGYSGLSVKVNSTTGPGHELRNAL  
WHTGNTPGQVRTLWHDPRHIGWKDFTAYRWRLSHRPKTGFIRVVMYEGKKIMADSGPIYD  
KTYAGGRLGLFVFSQEMVFFSCLKYECRDP

FIGURE 1D

# SEQUENCE- TSP-2

## nucleotide:

```

acggcatcca gtacagaggg gctggacttg gacccctgca gcagccctgc
acaggagaag cggcatataa agccgcgctg cccgggagcc gctcggccac
gtccaccgga gcatcctgca ctgcagggcc ggtctctcgc tccagcagag
cctgcgcctt tctgactcgg tccggaacac tgaaaccagt catcactgca
tcttttttggc aaaccaggag ctgagctgca ggaggcagga tggctctggag
gctggtcctg ctggctctgt ggggtgtggcc cagcacgcaa gctggtcacc
aggacaaaga cacgaccttc gaccttttca gtatcagcaa catcaaccgc
aagaccattg ggcgaagca gttccgcggg cccgaccccg gcgtgccggc
ttaccgcttc gtgcgctttg actacatccc accggtgaac gcagatgacc
tcagcaagat caccaagatc atgcggcaga aggagggtt cttcctcacg
gccagctca agcaggacgg caagtccagg ggcacgctgt tggctctgga
gggccccggg ctctcccaga ggcagttcga gatcgtctcc aacggccccg
cggacacgct ggatctcacc tactggattg acggcacccg gcatgtggtc
tccctggagg acgtcggcct ggctgactcg cagtgggaaga acgtcaccgt
gcaggtggct ggcgagacct acagcttgca cgtgggctgc gacctcatag
gaccagttgc tctggacgag cccttctacg agcacctgca ggcggaaaag
agccggatgt acgtggccaa aggctctgcc agagagagtc acttcagggg
tttgcttcag aacgtccacc tagtgtttga aaactctgtg gaagatattc
taagcaagaa ggggttgccag caaggccagg gagctgagat caacgccatc
agtgagaaca cagagacgct gcgcctgggt ccgcatgtca ccaccgagta
cgtgggcccc agctcggaga ggaggcccga ggtgtgcgaa cgctcgtgcg
aggagctggg aaacatggtc caggagctct cggggctcca cgctcctcgtg
aaccagctca gcgagaacct caagagagtg tcgaatgata accagtttct
ctgggagctc attggtggcc ctctaagac aaggaacatg tcagcttgct
ggcaggatgg ccggttcttt gcggaaaatg aaacgtgggt ggtggacagc
tgcaccacgt gtacctgcaa gaaatttaaa accatttgcc accaaatcac
ctgcccgcct gcaacctgcg ccagtccatc ctttggtgaa ggcgaatgct
gcccttcctg cctccactcg gtggacgggt aggagggtg gtctccgtgg
gcagagtgga ccagtgctc cgtgacgtgt ggctctggga ccagcagag
aggccgggtc tgtgacgtca ccagcaacac ctgcttgggg cctcogatcc
agacacgggc ttgcagtctg agcaagtgtg acaccgcgat ccggcaggac
ggcggctgga gccactggtc accttggctt tcatgctctg tgacctgtgg
agttggcaat atcacacgca tccgtctctg caactcccca gtgccccaga
tggggggcaa gaattgcaa gggagtggcc gggagaccaa agcctgccag
ggcgcgccat gcccaatcga tggccgctgg agcccctggg ccccggtggtc
ggcctgcact gtcacctgtg ccggtgggat ccgggagcgc acccgggtct
gcaacagccc tgagcctcag tacggaggga aggcctgcgt gggggatgtg
caggagcgtc agatgtgcaa caagaggagc tgccccgtgg atggctgttt
atccaacccc tgcttcccgg gagcccagtg cagcagcttc cccgatgggt
cctggtcatg cggcttctgc cctgtgggtt tcttgggcaa tggcaccac
tgtgaggacc tggacgagtg tgccctgggt cccgacatct gcttctccac
cagcaagggt cctcgtctgt tcaacactca gcctggcttc cactgcctgc
cctgcccgcc ccgatacaga gggaaaccag ccgtcggggg cggcctggaa
gcagccaaga cggaaaagca agtgtgtgag cccgaaaacc catgcaagga
caagacacac aactgccaca agcacgcgga gtgcatctac ctgggtcact
tcagcgaccc catgtacaag tgcgagtgcc agacaggcta cgcgggcgac
gggctcatct gcggggagga ctcggaacct gacggctggc ccaacctcaa
tctggtctgc gccaccaacg ccacctacca ctgcatcaag gataactgcc

```

FIGURE 2A

cccatctgcc	aaattctggg	caggaagact	ttgacaagga	cgggattggc
gatgcctgtg	atgatgacga	tgacaatgac	ggtgtgaccg	atgagaagga
caactgccag	ctcctcttca	atccccgcca	ggctgactat	gacaaggatg
aggttgggga	ccgctgtgac	aactgccctt	acgtgcacaa	ccttgcgccag
atcgacacag	acaacaatgg	agaggggtgac	gcctgctccg	tggacattga
tggggacgat	gtcttcaatg	aacgagacaa	ttgtccctac	gtctacaaca
ctgaccagag	ggacacggat	ggtgacgggtg	tgggggatca	ctgtgacaac
tgccccctgg	tgcacaaccc	tgaccagacc	gacgtggaca	atgaccttgt
tgyggaccag	tgtgacaaca	acgaggacat	agatgacgac	ggccaccaga
acaaccagga	caactgcccc	tacatctcca	acgccaacca	ggctgaccat
gacagagacg	gccagggcga	cgctgtgac	cctgatgatg	acaacgatgg
cgtccccgat	gacagggaca	actgccggct	tgtgttcaac	ccagaccagg
aggacttggg	cgggtgatgga	cggggtgata	tttgtaaaga	tgattttgac
aatgacaaca	tcccagatat	tgatgatgtg	tgctctgaaa	acaatgccat
cagtgaagaca	gacttcagga	acttccagat	ggtccccctg	gatcccaaag
ggaccaccca	aattgatccc	aactgggtca	ttcgccatca	aggcaaggag
ctggttcaga	cagccaactc	ggacccccgc	atcgctgtag	gttttgacga
gtttgggtct	gtggacttca	gtggcacatt	ctacgtaaac	actgaccggg
acgacgacta	tgctggcttc	gtctttggtt	accagtcaag	cagccgcttc
tatgtggtga	tgtggaagca	ggtgacgcag	acctaactgg	aggaccagcc
cacgcggggc	tatggctact	ccggcggtgc	cctcaagggtg	gtgaactcca
ccacggggac	gggcgagcac	ctgaggaacg	cgctgtggca	cacgggggaa
acgccggggc	aggtgcgaac	cttatggcac	gacccagga	acattggctg
gaaggactac	acggcctata	ggtggcacct	gactcacagg	ccaagaccg
gctacatcag	agtcttagtg	catgaaggaa	aacagggtcat	ggcagactca
ggacctatct	atgaccaaac	ctacgctggc	gggcggctgg	gtctatttgt
cttctctcaa	gaaatggtct	atttctcaga	cctcaagtac	gaatgcagag
atattttaa	aagatttgct	gcatttccgg	caatgccctg	tgcatgccat
ggtccctaga	cacctcagtt	cattgtgggtc	cttgcggtct	ctctctctag
cagcacctcc	tgtcccttga	ccttaactct	gatgggttct	cacctcctgc
cagcaacccc	aaacccaagt	gccttcagag	gataaatatc	aatggaaactc
agagatgaac	atctaaccga	ctagaggaaa	ccagtttggg	gatatatgag
acttttatgtg	gagtgaata	tgggcatgcc	attacattgc	ttttcttgt
ttgtttaaaa	agaatgacgt	ttacatataa	aatgtaatta	cttattgtat
ttatgtgtat	atggagttga	agggaatact	gtgcataagc	cattatgata
aattaagcat	gaaaaatatt	gctgaactac	ttttggtgct	taaagttgtc
actattcttg	aattagagtt	gctctacaat	gacacacaaa	tcccgtctaa
taaattataa	acaagggtca	attcaaattt	gaagtaatgt	tttagtaagg
agagattaga	agacaacagg	catagcaa	gacataagct	accgattaac
taatcggaac	atgtaaaaca	gttacaaaaa	taaacgaact	ctcctcttgt
cctacaatga	aagccctcat	gtgcagtaga	gatgcagttt	catcaaagaa
caaacatcct	tgcaaatggg	tgtgacgcgg	ttccagatgt	ggatttgga
aaacctcatt	taagtaaaag	gttagcagag	caaagtgcgg	tgcttttagct
gctgcttggtg	ccgttggtggc	gtcggggagg	ctcctgcctg	agcttccttc
cccagctttg	ctgcctgaga	ggaaccagag	cagacgcaca	ggccggaaaa
ggcgcatcta	acgcgtatct	aggctttggg	aactgcggac	aagttgcttt
tacctgattt	gatgatacat	ttcattaagg	ttccagttat	aaatattttg
ttaatattta	ttaagtgact	atagaatgca	actccattta	ccagtaactt
attttaaata	tgcttagtaa	cacatatgta	gtataatttc	tagaaacaaa
catctaataa	gtatataatc	ctgtgaaaat	atgaggcttg	ataatattag
gttgtcacga	tgaagcatgc	tagaagctgt	aacagaatac	atagagaata

FIGURE 2B

atgaggagtt tatgatggaa ccttaatatata taatgttgcc agcgatttta  
 gttcaatatt tggtactggt atctatctgc tgtatatgga attcttttaa  
 ttcaaacgct gaaaacgaat cagcatttag tcttgccagg cacaccaat  
 aatcagtcac gtgtaatatg cacaagtttg tttttgtttt tgttttttt  
 gttggttggt ttttttgctt taagttgcat gatctttctg caggaaatag  
 tcaactcatcc cactccacat aaggggttta gtaagagaag tctgtctgtc  
 tgatgatgga tagggggcaa atctttttcc cctttctggt aatagtcac  
 acatttctat gccaaacagg aacgatccat aacttttagtc ttaatgtaca  
 cattgcattt tgataaaaatt aattttgttg tttcctttga ggttgatcgt  
 tgtgttggtt tgctgcactt tttacttttt tgcgtgtgga gctgtattcc  
 cgagacaacg aagcgttggtg atacttcatt aaatgtagcg actgtcaaca  
 gcgtgcagggt tttctgtttc tgtgttggtg ggtcaaccgt acaatggtgt  
 gggaatgacg atgatgtgaa tatttagaat gtaccatatt ttttgtaa  
 tatttatggt tttctaaaca aatttatcgt ataggttgat gaaacgtcat  
 gtgttttgcc aaagactgta aatatttatt tatgtgttca catggtcaaa  
 atttcaccac tgaaaccctg cacttagcta gaacctcatt tttaaagatt  
 aacaacagga aataaattgt aaaaaagggt ttct

**protein:**

MVWRLVLLALWVWPSTQAGHQDKDTTFDLFSISNINRKTIGAKQFRGPDGPVPAYRF  
 VRFDYIPPVNADDLSKITKIMRQKEGFFLTAQLKQDGKSRGTLLEGLSQRQFE  
 IVSNGPADTLDLTYWIDGTRHVVSLEDVGLADSQWKNVTVQVAGETYSLHVGCDLIG  
 PVALDEPFYEHQAESRMYVAKGSARESHFRGLLQNVHLVFENSVEDILSKKGCQO  
 GQGAIEINAISENTETLRLGPHVTTEYVGPSSERRPEVCERSCEELGNMVQELSGLHV  
 LVNQLSENLRVSNDNQFLWELIGGPPKTRNMSACWQDGRFFAENETWVVDSTCTCT  
 CKKFKTICHQITCPPATCASPSFVEGECCPSCLHSVDGEEGWSPPAEWTQCSVTCGS  
 GTQQRGRSCDVTSNTCLGPSIQTRACSLSKCDTRIRQDGGWSHWSPWSSCSVTCGVG  
 NITRIRLCNSPVPQMGGKNCKGSGRETKACQGAPCPIDGRWSPWSPWSACTVTCAGG  
 IRERTRVCNSPEPQYGGKACVGDVQERQMCNKRSCPVDGCLSNPCFPGAQCSSFPDG  
 SWSCGFPCPVGFLGNGTHCEDLDECALVPDICEFSTSKVPRCVNTQPGFHCLPCPPRYR  
 GNQPVGVGLEAAKTEKQVCEPENPCKDKTHNCHKHAECIYLGHFSDPMYKCECQTGY  
 AGDGLICGEDSDLDGWPNNLVCATNATYHCICKDNCPHLPNSGQEDFDKDGIGDADC  
 DDDNDGVTDEKDNCQLLENPRQADYDKDEVGDRCDCNCPYVHNPAQIDTDNNGEGDA  
 CSVDIDGDDVFNERDNCPPYVNTDQRDTDGDGVGDHCDNCPLVHNPQDQDNDLVG  
 DQCDNNEDIDDDGHQNNQDNCPIISNANQADHHRDGGQDADCDPDDNDGVPDDRNC  
 RLVFNPDQEDLDGDGRGDICKDDFDNDNIPDIDDVCPENNAISETDFRNFQMVPLDP  
 KGTQIDPNWVIRHQKELVQTANSDPGIAVGFEFGSVDFSGTFYVNTDRDDDYAG  
 FVFGYQSSSRFYVVMWKQVTQTYWEDQPTRAYGYSGVSLKVVNSTTGTGEHLRNALW  
 HTGNTPGQVRTLWHDPRNIGWKDYTAIRWHLTHRPKTGYIRVLVHEGKQVMADSGPI  
 YDQTYAGGRLGLFVFSQEMVYFSDLKYECRDI

FIGURE 2C

# HT2143 Report

## RECORD INFORMATION

Gene ID: 2081  
Sequence ID: 2143  
Protein ID: 2125  
Sequence name: thrombospondin 4  
Genome: nucleus  
Taxon: Homo sapiens  
Locus: 2081  
Common Name: thrombospondin 4  
Role ID: 40

Coding sequence length: 2886 nt  
Transcript sequence length: 3074 nt  
Expression data: THC168897

## ACCESSION DATA

HT2143 is derived from accessions(s):

SP:P35443 (THROMBOSPONDIN 4 PRECURSOR.)  
GB:Z19585 (thrombospondin-4 (Homo sapiens))  
GB:Z19585 (H.sapiens mRNA for thrombospondin-4)  
PIR:A55710 (thrombospondin 4 precursor - human)

## cDNA FEATURES

Feature	End 5	End 3
coding_seq	28	2913
3'UT	2914	3074

## SEQUENCE

nucleotide:

gaattccggggagcaggaagagccaacatgctggccccgcgcggagccgcgctcctcctg  
ctgcacctggtcctgcagcgggtggctagcggcagggcggccaggccacccccagggtctt  
gaccttctcccatcttccagtcagaggctaaacccaggcgctctgctgccagtcctgaca  
gaccccgccctgaatgatctctatgtgatttccaccttcaagctgcagactaaaagttca  
gccaccatcttcgggtctttactcttcaactgacaacagtaaatattttgaatttactgtg  
atgggacgcttaagcaaagccatcctccgttacctgaagaacgatgggaaggtgcatttg

FIGURE 3A



gtgggttttcaacaacctgcagctggcagacggaaggcgccacaggatcctcctgaggctg  
agcaatttgcagcgagggggccggctccctagagctctacctggactgcacccaggtggat  
tcogttcacaatctccccagggcctttgctggccccctccagaaacctgagaccattgaa  
ttgaggactttccagaggaagccacaggacttcttggaaagagctgaagctgggtgggtgaga  
ggctcaactgttccaggtggccagcctgcaagactgcttctgcagcagagtgagccactg  
gctgccacagggcacaggggactttaaccggcagttcttgggtcaaatgacacaattaaac  
caactcctgggagaggtgaaggacottctgagacagcagggttaaggaaacatcatttttg  
cgaaacaccatagctgaatgccaggcttgcgggtcctctcaagtttcagtctccgacccca  
agcacgggtgggtgcggccggctccccctgcacggcccaacacggccacctcgtcgggtgtgac  
tccaacccatgtttccgaggtgtccaatgtaccgacagtagagatggcttccagtggtggg  
ccctgccccgaggggtacacaggaaacgggatcacctgtattgatgttgatgagtgcaaa  
taccatccctgctacccgggctgcaactgcataaatttgtctcctggcttcagatgtgac  
gctgcccagtggggtctcacaggccccatgggtgcagggtgttgggatcagttttgccaag  
tcaaacaaagcaggtctgcactgcattgatgagtgctgaaatggagcgtgcgttcccaac  
tcgatctgcgttaatactttgggatcttaccgctgtgggccttgtaagccggggtatact  
gggtgatcagataaggggatgcaagtggaagaaactgcagaaacccagagctgaaccct  
tgcagtgatgaatgccagtgcatggaagagaggcagggggatgtgacatgtgtgtgtgga  
gtcgggttgggctggagatggctatatctgtggaaaggatgtggacatcgacagttacccc  
gacgaagaactgccatgctctgccagggaactgtaaaaaggacaactgcaaatatgtgcca  
aattctggccaagaagatgcagacagagatggcattggcgacgcttgtgacgaggatgct  
gacggagatgggatcctgaatgagcaggataactgtgtcctgattcataantgtggacca  
aggaacagcgataaagatatctttggggatgcctgtgataactgcctgagtgctttaa  
aacgaccagaaaagacaccgatggggatgggaagaggagatgcctgtgatgatgacatggat  
ggagatgggaataaaaaacattctgggacaactgccccaaaatttcccaatcgtgaccaacgg  
gacaaggatgggtgatgggtgtgggggatgcctgtgacagttgtcctgatgtcagcaaccct  
naccagctctgatgtggataatgatctggttggggactcctgtgacaccaatcaggacagt  
gatggagatgggcaccaggacagcacagacaactgccccacggctattaacagtggccag  
ctggacaccgataaaggatgggaattgggtgacgagtgatgatgatgatgacaatgatgg  
atcccagacctgggtgccccctggaccagacaactgcccggctgggtccccaacccagcccag  
gaggatagcaacagcgacggagtgaggagacatctgtgagtgctgactttgaccaggaccag  
gtcatcgatcggatcgacgctctgccagagaacgcagagggtcacctgaccgacttcagg  
gcttaccagaccgtgggctggatcctgaaggggatgccagatcgatcccaactgggtg  
gtcctgaaccaggggcatggagattgtacagaccatgaacagtgatcctggcctggcagtg  
gggtacacagcttttaatggagttgacttcgaagggaccttccatgtgaataccagaca  
gatgatgactatgcaggctttatctttgggtaccaagatagctccagcttctacgtggtc  
atgtggaagcagacggagcagacatattgggaagccacccattccgagcagttgcagaa  
cctggcattcagctcaaggtctgtgaagtctaagacagggtccaggggagcatctccggaac  
tccctgtggcacacgggggacaccagtgaccagggtcaggctgctgtggaaggactccagg  
aatgtgggctgggaaggacaagggtgtcctaccgctgggttccctacagcacaggccccagg  
ggctacatcagggtacgattttatgaaggctctgagttgggtggctgactctggcgtcacc  
atagacaccacaatgcgtggaggccgacttggcgttttctgcttctctcaagaaaacatc  
atctgggtccaaacctcaagtatcgctgcaatgacaccatccctgaggacttccaagagttt  
caaaccacagaatttcgaccgcttcgataattaaaccaagggaagcaatctgtaactgcttt  
tcggaacactaaaaccatataatatttaacttcaattttcttttagcttttaccacccaa  
atatatcaaaaacgttttatgtgaatgtggcaataaaggagaagagatcatttttataaaaa  
aaaaaaaaaaaaaa

# protein:

MLAPRGAAVLLLHLVLQRWLAAGAQA TPQVFDLLPSSSQRLNPGALLPVL TDPALNOLYV  
ISTFKLQTKSSATIFGLYSSTDNSKYF EFTVMGRLSKAILRYLKNDGKVHLVVFNNLQLA  
DGRHRIRILLRLSNLQRGAGSLELYLD CIOVDSVHNLPRAFAGPSQKPE TIELRTFQRKPQ  
DFLEELKLVVRGSLFQVASLQDCFLQ QSEPLAATGTGDFNRQFLGQMTQLNQLLG EVKDL  
LRQVKETSFLRNTIAECQACGPLKFS PTSTV VAPAP PAPPTRPPRRCD SNPCFRGVQ  
CTDSRDGFQCGPCPEGYTGNGITC IDVDECKYHPCYPGVHCINLS PGFRCDACFVGTGP  
MVQGVGISFAKSNKQVCTDIDECRNGAC VPNSICVNTLGSYRCGPCKPGYTGDQIRGCKV  
ERNCRNP ELPNFC SVNAQCIEROGDVT CVCVGWAGDGYICGKDVIDSY PDEELPC SAR  
NCKKDNCKYVPNSGQEDADRDGIGD ADEADGDGILNEQONCVLIHNV DQRNSDKDIFG  
DACDNCLSVLNNQKDTDCDGRGDACDD DMDGDGIKNILDNCPKFPNRDQRDKDGDGVGD

FIGURE 3B

[illegible][illegible]

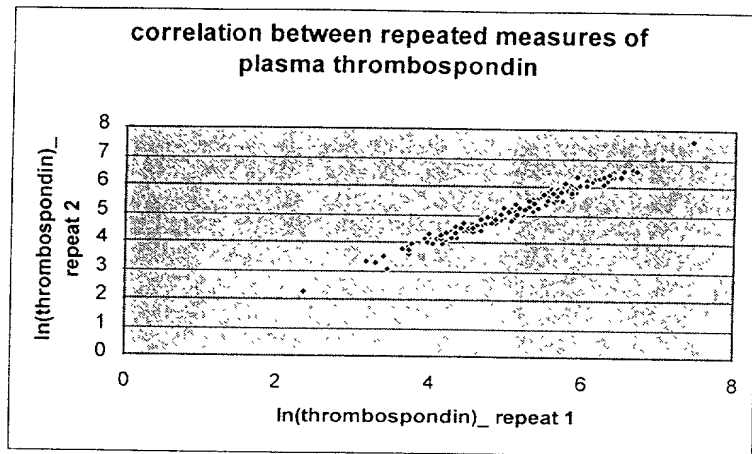


FIGURE 4